

SEQUENCE LISTING

<110> Tchaga, Grigory S.
Jokhadze, George

<120> Metal Ion Affinity Tags and Methods for
Using the Same

<130> CLON-056CIP

<140> US 09/858,332

<141> 2001-05-15

<150> 09/404,017

<151> 1999-09-23

<150> 60/101,867

<151> 1998-09-25

<160> 23

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 1

His Leu Ile His Asn Val His Lys Glu Glu His Ala His Ala His Asn
1 5 10 15

<210> 2

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 2

His Asp Asp His Asp Asp His Asp Asp His Asp Asp His
1 5 10 15
Asp Asp

<210> 3

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 3

His Glu Glu His Glu Glu His Glu Glu His Glu Glu His
1 5 10 15

<210> 4

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 4

His Asp Glu His Asp Glu His Asp Glu His Asp Glu His
1 5 10 15
Asp Glu

<210> 5

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 5

His Glu Asp His Glu Asp His Glu Asp His Glu Asp His
1 5 10 15
Glu Asp

<210> 6

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> enterokinase cleavage site

<400> 6

Ile Glu Gly Arg
1

<210> 7

<211> 6

<212> PRT

<213> Artificial Sequence

<220>
<223> a factor Xa cleavage site

<400> 7
Leu Val Pro Arg Gly Ser
1 5

<210> 8
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> a thrombin cleavage site

<400> 8
His Pro Phe His Leu Val Ile His
1 5

<210> 9
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> a renin cleavage site

<400> 9
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5 10

<210> 10
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> an immunological tag

<400> 10
Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 11
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> an immunological tag

<400> 11

Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> 12
<211> 11
<212> PR1
<213> Artificial Sequence

<220>
<223> an immunological tag

<400> 12
Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> 13
<211> 3430
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA sequence of vector containing cDNA of
recombinant enterokinase

<400> 13
gacgaaaggg cctcgtgata cgcctatatt tataggctaa tgcctatgata ataattggtt 60
cttagacgtc aggtggcact ttccggggaa atgtgcgggg aaccttatt tgtttatatt 120
tctaaatata ttcaaatatg tatccgctca tgagacata accctgataa atgcttcaat 180
aatattgaaa aaggaagagt atgagtattc aacatttcct tgcgcctt attcctttt 240
ttgcggcatt ttgccttctt gtttttctc acccagaaac gctgggtgaaa gtaaaagatg 300
ctgaagatca gttgggtgca cgaagtgggtt acatcgaaat ggatctcaac agcggtaaga 360
tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gaggactttt aaagtctctg 420
tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcanctcggg cgcgcgatac 480
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg 540
gcattgacagt aagagaatta tgcagtgcctg ccataacctt gaggatgata actgcggcca 600
acttacttct gacaacgacg ggaggaccca aggagctaac cgtttttttg cacaacatgg 660
gggatcctgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaaag 720
acgagcgtga caccacgatg cctgtagcaa tggcaacaa gttgcgcaaa ctattaactg 780
gcgaactact tactctagct tcccggaac aattaataga ctggatggag gcggataaag 840
ttgcaggacc acttctgcgc tcggcccttc cggctggctg qtttattgct gataaatctg 900
gagccggtga gcgtgggtct cgcggtatca ttgcagcaat ggggccaagt ggtaagccct 960
cccgatcgt agttatctac acgacgggga gtcaggcaac tatggatga cgaactagac 1020
agatcgtga gataggtgcc tcaactgatta agcattggtt actgtcagac caagttaact 1080
catatatact ttgattgat ttaaaacttc atttttaatt tanaaggatc taggtgaga 1140
tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgtc cactgagcgt 1200
cagacccctt agaaaagatc aaaggntctt cttgagatcc tttttttctg cgcgtaactt 1260
gctgcttgca aacaaanaaa ccaccgctac cagcgttggg ttgtttgccc gatcaagagc 1320
taccactct tttccgaag gtaactggtt tcagcagagc gcagatacca aatactgtcc 1380
ttctagtgtg gccgtagtta ggcaccact tcaagaactc tctagcaccg cctacataac 1440
tcgctctgct aatcctgtta ccagtggctg ctgcagtggt cgataagtcg tgtcttaccg 1500
ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtc 1560
cgtgcacaga gccagcttg gagcgaacga cctacaccca actgagatcc ctacagcgtg 1620
agctatgaga aagcgcacag cttcccgag ggagaaagcg ggacaggtat ccgtaagcgt 1680
gbgcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct 1740
ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 1800

```

aggggggcgg agcctatgga aaaacgccag caacggggcc tttttacggt tcttggcctt 1860
ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggntaacgg 1920
tattaccgcc tttgagtggg ctgataccgc tcgccgcagg cgaacgacgg agcgcagcga 1980
gtcagtgagc gaggaagcgg aagagcgccc aatacgcnaa ccgcctctcc ccgcgcgttg 2040
gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcaqtgagcg 2100
caacgcaatt aatgtgagtt agtcactca ttaggcaccc caggetttac actttatgct 2160
tccggtcgt atgtgtgtg gaattgtgag cggataacaa ttccacacag gaacagctt 2220
tgaccatgat tacgccaagc ttgaaggatc atctctccaa caatgtccac aaagaggagc 2280
acgctcatgc ccacaacaag atcgatattg tcggaggagc tgactccaga gaaggagcct 2340
ggccttgggt cgttgcctctg tatttcgacg atcaacaggt cgcggagct tctctgggtg 2400
gcagggattg gctgggtgctg gccgccact gcgtgtacgg gngaaatag gagcctcta 2460
agtggaaagc agtgctaggc ctgcataatg catcaatct gacttctct cagatagaaa 2520
ctagggtgat tgaccaaatt gtcataaacc cacactacaa taacgggaga aagaacaatg 2580
acattgccat gatgcattt gaaatgaaag tgaactacac agattatata cagcctatct 2640
gtttaccaga agaaaatcaa gtttttcccc caggaagaaat ttgttctatt gctggctggg 2700
gggcacttat atatcaaggc tctactgcag acgtactgon agaagctgac gttccccctt 2760
talcaaatga gaaatgtcaa caacagatgc cagaalatna cattacggaa aatatgggtg 2820
gtgcaggcta tgaagcagga ggggtagatt cttgtcaggg ggaltcaggg ggaccactca 2880
tgtgccaaaga naacaacaga tggctcctgg ctggcglgac gtcatttgga tatcaatgtg 2940
cactgcctaa tcgcccgagg gtgtatgcc gggtcccaag gttcacagag tggatacaaa 3000
gttttctaca tgagctcgta attagctgag aattcactgg ccgtcgtttt acaacgtcgt 3060
gagcctggga aaacctggc gttacccaac ccgatcgccc tcccccaacag ttgcgcagcc 3120
ccagctggcg taatagcgaa gaggcccgca ttctctctac gcatctgtgc ggtatttcc 3180
tgaatggcga atggcgctg atcggttatt ttctctctac gcatctgtgc ggtatttcc 3240
accgcataat gtgcactctc agtacaatct gctctgltgc cgcatagtta agccagcccc 3300
gacacccgcc aacaccgct gacgcgcctt gacgggcttg tctgtccccc gcatecgctt 3360
acagacaagc tgtgaccgtc tcggggagct gcatgtgtca gaggttttca ccgtcatcnc 3420
cgaaacgcgc

```

<210> 14

<211> 279

<212> PRT

<213> Artificial Sequence

<220>

<223> protein sequence of vector containing cDNA of recombinant enterokinase

<400> 14

```

Met Thr Met Ile Thr Pro Ser Leu Lys Asp His Leu Ile His Asn Val
1      5      10      15
His Lys Glu Glu His Ala His Ala His Asn Lys Ile Asp Ile Val Gly
20      25      30
Gly Ser Asp Ser Arg Glu Gly Ala Trp Pro Trp Val Val Ala Leu Tyr
35      40      45
Phe Asp Asp Gln Gln Val Cys Gly Ala Ser Leu Val Ser Arg Asp Trp
50      55      60
Leu Val Ser Ala Ala His Cys Val Tyr Gly Arg Asn Met Glu Pro Ser
65      70      75      80
Lys Trp Lys Ala Val Leu Gly Leu His Met Ala Ser Asn Leu Thr Ser
85      90      95
Pro Gln Ile Glu Thr Arg Leu Ile Asp Gln Ile Val Ile Asn Pro His
100     105     110
Tyr Asn Lys Arg Arg Lys Asn Asn Asp Ile Ala Met Met His Leu Glu
115     120     125
Met Lys Val Asn Tyr Thr Asp Tyr Ile Gln Pro Ile Cys Leu Pro Glu
130     135     140

```

Glu Asn Gln Val Phe Pro Pro Gly Arg Ile Cys Ser Ile Ala Gly Trp
 145 150 155 160
 Gly Ala Leu Ile Tyr Gln Gly Ser Thr Ala Asp Val Leu Gln Glu Ala
 165 170 175
 Asp Val Pro Leu Leu Ser Asn Glu Lys Cys Gln Gln Met Pro Glu
 180 185 190
 Tyr Asn Ile Thr Glu Asn Met Val Cys Ala Gly Tyr Glu Ala Gly Gly
 195 200 205
 Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Glu
 210 215 220
 Asn Asn Arg Trp Leu Leu Ala Gly Val Thr Ser Phe Gly Tyr Gln Cys
 225 230 235 240
 Ala Leu Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Pro Arg Phe Thr
 245 250 255
 Glu Trp Ile Gln Ser Phe Leu His Glu Leu Val Ile Ser Glu Phe Thr
 260 265 270
 Gly Arg Arg Phe Thr Thr Ser
 275

<210> 15
 <211> 12
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> an amino acid sequence embodiment of the affinity
 purification site

<400> 15
 His Asn His Asn His Asn His Asn His Asn
 1 5 10

<210> 16
 <211> 48
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> a DNA sequence embodiment of the affinity
 purification site

<400> 16
 catctcatcc acaatgtcca caaagaggag cagcgtcatg cccacaac

48

<210> 17
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> a DNA sequence embodiment of the affinity
 purification site

<400> 17
 cataaccata accataacca taaccataac cataac

36

<210> 18
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> a DNA sequence embodiment of the affinity
purification site

<400> 18
catgatgatc atgatgatca tgatgatcat gatgatcatg atgatcatga lgat 54

<210> 19
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> a DNA sequence embodiment of the affinity
purification site

<400> 19
catgaggagc atgaggagca tgaggagcat gaggagcatg aggagcatga ggag 54

<210> 20
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> a DNA sequence embodiment of the affinity
purification site

<400> 20
catgatgagc atgatgagca tgagaaccat gagaaccatg aggatcatga ggat 54

<210> 21
<211> 9
<212> PRT
<213> Human

<220>
<221> VARIANT
<222> 9
<223> Xaa at position 9 is an amino acid with an
aliphatic or amide side chain.

<221> VARIANT
<222> 2
<223> Xaa at position 2 is an amino acid with an
aliphatic or amide side chain

<221> VARIANT
<222> 3
<223> Xaa at position 3 is an amino acid with an
aliphatic or amide side chain.

<221> VARIANT

<222> 5

<223> Xaa at position 5 is an amino acid with a basic side chain (except HIS) or an acidic side chain.

<221> VARIANT

<222> 6

<223> Xaa at position 6 is an amino acid with a basic side chain (except HIS) or an acidic side chain.

<221> VARIANT

<222> 7

<223> Xaa at position 7 is an amino acid with a basic side chain (except HIS) or an acidic side chain.

<400> 21

His Xaa Xaa His Xaa Xaa Xaa His Xaa
1 5

<210> 22

<211> 6

<212> PRT

<213> human

<400> 22

His Arg His Arg His Arg
1 5

<210> 23

<211> 9

<212> PRT

<213> human

<220>

<221> VARIANT

<222> 2

<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT

<222> 3

<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT

<222> 5

<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT

<222> 6

<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT

<222> 8

<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT

<222> 9

<223> Xaa .. an amino acid having an acidic side chain.

<400> 23

His Xaa Xaa His Xaa Xaa His Xaa Xaa

1

5